Many time series of practical interest show strong dependence (or long memory). This phenomenon appears visually as fractal behaviour and is connected to heavy tails of a distribution of the series. This aspect occurs for example in computer networks. It is empirically observed that network data follow fractal behaviour.

**What is the goal?**

The goal of this project is to describe statistical behaviour of network traffic models and estimate relevant parameters such as the memory parameter.

**Results**

We analysed a simulated network traffic model. We computed the estimator of the memory parameter for that simulated traffic model. The results indicate a good performance of the R/S estimator. Our simulations studies motivate further research on theoretical properties of the estimator and its precise relation to the optimal capacity of the network.

**Few definitions:**

- **Long memory**: A time series is said to have a long memory if values from the distant past still have an effect on present values.
- **Fractal behaviour**: It is a type of pattern used in technical analysis to predict a reversal in the current trend.
- **Computer network**: It can be defined as a group of computers and other devices connected in some ways in order to exchange data. It usually consists of many sources.
- **Heavy-tailed distribution**: It is a distribution that assigns relatively high probabilities to regions far from the mean or the median.
- **Long-range dependency**: It is a notion that can be defined as a slow decay of correlations. It is a phenomenon that may arise in the analysis of time series data.

**RESOURCES**

*Statistical analysis of non-standard time series models*

*Model description*

**Traffic Process from one source:**

\[
W(t) = \sum_{j \geq 1} (t_j \leq t \leq t_j + X_j), \quad t \in [0, \infty)
\]

**X_j random variables** = duration times of each transmission

**T_j points from Poisson Process** = beginning of a transmission

**Assumption 1:**

Duration follow heavy-tailed distribution:

\[
P(X > x) \sim const \cdot x^{- \alpha}, \quad \alpha \in (1,2)
\]

The parameter \(\alpha\) is related to the size

**Assumption 2:**

Sequences \(T_j\) and \(X_j\) are independent form each other.

**Central limit theorem not valid anymore for the long memory traffic models**

**Rescaled Range method**

\[
\hat{d}_{R/S} = \frac{\log(R_n/S_n)}{\log n} - \frac{1}{2}
\]

is the R/S estimator of the memory parameter \(d\) which is needed to construct a network in an optimal way. Intuitively, the bigger \(d\) is, the bigger the capacity of the network will be.

**Estimation of the memory parameter:**

\(d+1/2\) can be interpreted as the slope of a regression line of \(\log(R_n/S_n)\) against \(\log n\).

**Traffic models that exhibit long memory**

**Few definitions:**

- **Long memory**: A time series is said to have a long memory if values from the distant past still have an effect on present values.
- **Fractal behaviour**: It is a type of pattern used in technical analysis to predict a reversal in the current trend.
- **Computer network**: It can be defined as a group of computers and other devices connected in some ways in order to exchange data. It usually consists of many sources.
- **Heavy-tailed distribution**: It is a distribution that assigns relatively high probabilities to regions far from the mean or the median.
- **Long-range dependency**: It is a notion that can be defined as a slow decay of correlations. It is a phenomenon that may arise in the analysis of time series data.

**RESULTS**

We analysed a simulated network traffic model. We computed the estimator of the memory parameter for that simulated traffic model. The results indicate a good performance of the R/S estimator. Our simulations studies motivate further research on theoretical properties of the estimator and its precise relation to the optimal capacity of the network.

We also simulated a traffic model from \(M = 100\) sources. We chose \(\alpha = 1.3\). The simulated model is displayed on figure 1. The result of the regression procedure is shown on figure 2. The estimated parameter \(d\) is \(d = 0.350714\), whereas the formula \(d = 1 - \alpha/2\) gives 0.35. Hence, our estimation is very accurate.

**Technical details:**

- **Traffic Process from one source:**

\[
W(t) = \sum_{j \geq 1} (t_j \leq t \leq t_j + X_j), \quad t \in [0, \infty)
\]

\(W(t)\) is the total transmission from the beginning of the time series to the time \(t\).

\(X_j\) are the duration times of each transmission.

**Rescaled Range method**

\[
\hat{d}_{R/S} = \frac{\log(R_n/S_n)}{\log n} - \frac{1}{2}
\]

is the R/S estimator of the memory parameter \(d\) which is needed to construct a network in an optimal way. Intuitively, the bigger \(d\) is, the bigger the capacity of the network will be.

**Estimation of the memory parameter:**

\(d+1/2\) can be interpreted as the slope of a regression line of \(\log(R_n/S_n)\) against \(\log n\).

**Traffic models that exhibit long memory**

**Few definitions:**

- **Long memory**: A time series is said to have a long memory if values from the distant past still have an effect on present values.
- **Fractal behaviour**: It is a type of pattern used in technical analysis to predict a reversal in the current trend.
- **Computer network**: It can be defined as a group of computers and other devices connected in some ways in order to exchange data. It usually consists of many sources.
- **Heavy-tailed distribution**: It is a distribution that assigns relatively high probabilities to regions far from the mean or the median.
- **Long-range dependency**: It is a notion that can be defined as a slow decay of correlations. It is a phenomenon that may arise in the analysis of time series data.

**RESULTS**

We analysed a simulated network traffic model. We computed the estimator of the memory parameter for that simulated traffic model. The results indicate a good performance of the R/S estimator. Our simulations studies motivate further research on theoretical properties of the estimator and its precise relation to the optimal capacity of the network.

We also simulated a traffic model from \(M = 100\) sources. We chose \(\alpha = 1.3\). The simulated model is displayed on figure 1. The result of the regression procedure is shown on figure 2. The estimated parameter \(d\) is \(d = 0.350714\), whereas the formula \(d = 1 - \alpha/2\) gives 0.35. Hence, our estimation is very accurate.

**Technical details:**

- **Traffic Process from one source:**

\[
W(t) = \sum_{j \geq 1} (t_j \leq t \leq t_j + X_j), \quad t \in [0, \infty)
\]

\(W(t)\) is the total transmission from the beginning of the time series to the time \(t\).

\(X_j\) are the duration times of each transmission.

**Rescaled Range method**

\[
\hat{d}_{R/S} = \frac{\log(R_n/S_n)}{\log n} - \frac{1}{2}
\]

is the R/S estimator of the memory parameter \(d\) which is needed to construct a network in an optimal way. Intuitively, the bigger \(d\) is, the bigger the capacity of the network will be.

**Estimation of the memory parameter:**

\(d+1/2\) can be interpreted as the slope of a regression line of \(\log(R_n/S_n)\) against \(\log n\).

**Traffic models that exhibit long memory**

**Few definitions:**

- **Long memory**: A time series is said to have a long memory if values from the distant past still have an effect on present values.
- **Fractal behaviour**: It is a type of pattern used in technical analysis to predict a reversal in the current trend.
- **Computer network**: It can be defined as a group of computers and other devices connected in some ways in order to exchange data. It usually consists of many sources.
- **Heavy-tailed distribution**: It is a distribution that assigns relatively high probabilities to regions far from the mean or the median.
- **Long-range dependency**: It is a notion that can be defined as a slow decay of correlations. It is a phenomenon that may arise in the analysis of time series data.

**RESULTS**

We analysed a simulated network traffic model. We computed the estimator of the memory parameter for that simulated traffic model. The results indicate a good performance of the R/S estimator. Our simulations studies motivate further research on theoretical properties of the estimator and its precise relation to the optimal capacity of the network.

We also simulated a traffic model from \(M = 100\) sources. We chose \(\alpha = 1.3\). The simulated model is displayed on figure 1. The result of the regression procedure is shown on figure 2. The estimated parameter \(d\) is \(d = 0.350714\), whereas the formula \(d = 1 - \alpha/2\) gives 0.35. Hence, our estimation is very accurate.